A High-Level Object-Oriented Model for Representing Relationships in an Electronic Medical Record

Robert H. Dolin, MD Southern California Kaiser-Permanente 3111 W. Orange Av, Anaheim, CA 92804

ABSTRACT

The importance of electronic medical records to improve the quality and cost-effectiveness of medical care continues to be realized. This growing importance has spawned efforts at defining the structure and content of medical data, which is heterogeneous, highly interrelated, and complex. Computer-assisted data modeling tools have greatly facilitated the process of representing medical data, however the complex inter-relationships of medical information can result in data models that are large and cumbersome to manipulate and view. This report presents a high-level object-oriented model for representing the relationships between objects or entities that might exist in an electronic medical record. By defining the relationship between objects at a high level and providing for inheritance, this model enables relating any medical entity to any other medical entity, even though the relationships were not directly specified or known during data model design.

INTRODUCTION

There is a growing need to define, standardize, and model the complex informational needs of medicine. In 1991, the Institute of Medicine emphasized the need for computer-based patient records to improve the quality and cost-effectiveness of patient care [1]. In 1993, the American Hospital Association noted the importance of computerized data to evaluate the likely outcomes of alternative treatment options [2]. President William J. Clinton's American Health Security Act stresses the need for information on health outcomes and calls for national standards for information systems, clinical data, and minimal health data sets [3].

Computer-assisted data modeling tools and techniques are being used to help define the structure and content of medical data. Entity-Relationship (E-R) modeling [4] and Object-Oriented Analysis (OOA) [5] are two techniques in common use [6-9]. In E-R modeling, entities (such as ENCOUNTERS or PATIENTS) have attributes (such as SEX or ADDRESS) and relationships to one another (such as PATIENTS having MANY ENCOUNTERS). E-R models map directly into relational databases. In OOA, the E-R model is extended such that entities are referred to as objects, and inheritance is supported. Using inheritance, an object can inherit the attributes of another object. For example, if the object PERSON has attributes SEX and ADDRESS, and if the object PATIENT is a specialization of PERSON, then PATIENT inherits the attributes SEX and ADDRESS from PERSON. Object-oriented data models can map to object-oriented databases or to relational databases. The concept of the E-R relational 'primary key' remains valid in OOA, meaning that both E-R and OOA require a mechanism of identifying unique instances of entities or objects.

E-R and OOA tools graphically depict the relationship between any two entities or objects as a line connecting them. These graphical data model depictions become complex as the number of objects increases, particularly when there is a large number of relationship lines. An average data model contains 35 objects, a large model contains 110, and for domains with several sub-domains there may be 200 to 500 objects [5]. If all objects related to all other objects, the number of relationships would be calculable using the formula n(n-1)/2, where n equals the number of objects. The Message Standards Developers Subcommittee of the American National Standards Institute **Healthcare Informatics Standards Planning** Panel (ANSI-HISPP) has developed a rough-cut high-level data model of the healthcare domain wherein approximately fifty medical objects are discussed [7]. If each of these objects related to each other, the number of relationships would equal n(n-1)/2 = 50(49)/2 = 1225. In addition, objects of a single class may relate to one another, there may be more than one relationship between two objects, and normalization of each many-to-many relationship potentially results in a new object and relationship.

One approach to managing all these interrelationships is to find an attribute that is common to all objects or entities. A consideration for this common attribute is TIME, meaning that two objects occurring at the same TIME for the same patient are assumed to relate to one another. For example, a patient having an ENCOUNTER at one TIME also has a PROCEDURE performed at the same TIME. One can infer that the PROCEDURE occurred during the ENCOUNTER. This solution will work in many situations, but in others, a precise relationship must be made and documented. A lab test result that was available during the time a patient was being seen does not necessarily indicate that the provider saw the result.

Obviously, every object will not need to relate to every other object. However there are a great number of required and potentially important relationships. The relationship of SYMPTOMS to PROBLEMS is necessary to support Weed's Problem-Oriented Medical Record [10]. Outcome analysis requires knowing the SERVICES that were provided for each PROBLEM [11,12]. Healthcare reform requires knowing the relationship between ORDERS and their INDICATIONS. When automated decision support is used, a provider may document the ALGORITHM used to arrive at a DIAGNOSIS or TREATMENT PLAN. Other relationships of potential interest include SYMPTOMS or RESULTS that support a DIAGNOSIS; ENCOUNTERS that address a certain PROBLEM: ORDERS generated during an ENCOUNTER; MEDICATIONS administered during a SERVICE or prescribed during an ENCOUNTER; and more.

Rather then attempting to exhaustively identify and define every potential relationship that might exist in an electronic medical record, the model described in this report provides a highlevel mechanism for relating any two objects, whether they be instances of the same object, or two different objects.

DESCRIPTION OF MODEL

Figure 1 gives a brief review of the modeling techniques used in this report. The graphical representation of objects, entities and relationships are not fully standardized. The Medical Informatics Technical Committee of the European Committee for Standardization (CEN / TC251) represents object relationships slightly different in two of its preliminary standards documents [6,8]. United States medical informatics standards are commonly modeled using the representation described in [5]. This nomenclature is used by the ANSI-HISPP in describing its high-level data model of the healthcare domain [7] and by the Institute of Electrical and Electronic Engineers (IEEE) medical device communications standards [9]. The model described here will follow the representation described in [5].

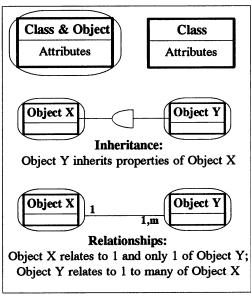


Figure 1. Object-Oriented Modeling Representation.

The high-level object-oriented data model for representing relationships among the objects or entities in an electronic medical record is shown in Figure 2. The object MEDICAL ENTITY OCCURRENCE represents the occurrence of a

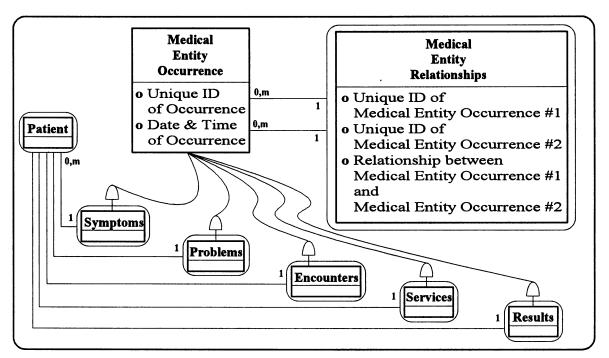


Figure 2. A high-level object-oriented data model for specifying relationships among objects in an electronic medical record.

medical entity at a specific date and time. Medical entities include SYMPTOMS. PROBLEMS, ENCOUNTERS, SERVICES, RESULTS, and more. They inherit attributes from the object class MEDICAL ENTITY OCCURRENCE, thus each medical entity contains the date and time of occurrence. **Examples of MEDICAL ENTITY** OCCURRENCES include the SERVICE of a Flexible Sigmoidoscopy on February 3, 1994; the SYMPTOM of Chest Pain that occurs on December 11, 1993; or the RESULT of a serum sodium value on March 24, 1994. The relationship between PATIENT and the medical entities SYMPTOMS, PROBLEMS, ENCOUNTERS, SERVICES, and RESULTS is zero-to-many on the PATIENT side, meaning a PATIENT can have at least zero and at most many instances of the entity occurrence. The relationship between each medical entity and PATIENT is one, meaning that each occurrence is specific to one and only one PATIENT.

MEDICAL ENTITY OCCURRENCES are associated with one another in a many-to-many relationship, since any entity occurrence may need to be related to one or more other entity occurrences. Because attributes are necessary to fully describe the relationship between two

MEDICAL ENTITY OCCURRENCES, the many-to-many relationship is normalized resulting in the creation of the object MEDICAL ENTITY RELATIONSHIPS. This object has an attribute 'Relationship Between Medical Entity Occurrence #1 and Medical Entity Occurrence #2' which is used to describe the nature of the relationship between two MEDICAL ENTITY OCCURRENCES, such as SYMPTOMS that SUPPORT THE DIAGNOSIS OF a particular DIAGNOSIS, or ORDERS that ARE ORDERED BECAUSE OF particular PROBLEMS.

Note further in Figure 2, the attribute 'Unique ID of Occurrence' of object MEDICAL ENTITY OCCURRENCE. Several authors note the potential pitfalls of trying to determine a unique object identifier during data model design, and feel it is preferable to defer its consideration to the implementation phase [5,7]. While the model in Figure 2 does show a unique identifier for medical entities, it does not attempt to fully define what that identifier should be. But the model does include the unique identifier as an inheritable attribute. In this way, all specializations of MEDICAL ENTITY OCCURRENCE, such as SYMPTOMS and PROBLEMS, will have their own unique

identifier structured in the same way. If each object structured its unique identifier in a different way, it would be difficult or impossible for a single table to store relationships between any two objects. But by providing for the inheritance of a common unique identifier, this model guarantees that all unique identifiers will have a common structure, thus enabling all specializations of MEDICAL ENTITY OCCURRENCE to be related to one another in MEDICAL ENTITY RELATIONSHIPS.

The unique identifier of an object is generally system dependent and implementation specific, therefore only general strategies can be described here. One approach is to use a standardized medical coding scheme, such as SNOMED International [13], to codify all medical entities. Then, a unique identifier might equal: Patient ID + SNOMED Code + Date & Time of occurrence. However, in a fully implemented electronic medical record, multiple providers may describe the same entity as it occurred at the same date and time. In fact, a single provider may describe the same entity herself again at a later date. Thus, the unique identifier may need to include Provider ID and Date & Time entered. Another approach is to allow the system to assign a unique sequence number to each entity occurrence, in which case the unique identifier of Patient ID + SNOMED Code + Sequence Number may suffice. However, if a given SNOMED code can be recorded in more then one table, the unique identifier may need to include the table name. In this case the unique identifier might simply be: Table Name + Sequence Number.

The following scenario illustrates how the proposed data model allows any object to relate to any other object. A patient seen in clinic today reports a symptom of chest pain that occurred three days ago. An electrocardiogram is performed and is normal. The provider orders a treadmill test. The clinic visit is an ENCOUNTER; chest pain is a SYMPTOM; the normal ECG is a RESULT; the treadmill test is an ORDER. Any of these MEDICAL ENTITY OCCURRENCES can be related to any other in MEDICAL ENTITY RELATIONSHIPS. This is shown in Table 1.

Table 1. Relating Medical Entity Occurrences.

ENTITY1 ID*	ENTITY2 ID*	RELATIONSHIP
Clinic Visit	Chest Pain	Elicited During
Clinic Visit	ECG	Ordered During
Clinic Visit	ECG	Reviewed During
Clinic Visit	Treadmill	Ordered During
ECG	Chest Pain	Reason for Order
Treadmill	Chest Pain	Reason for Order

^{*} The actual values stored are the unique identifiers.

DISCUSSION

The preceding discussion has shown how any medical entity can be related to any other medical entity, with provisions to record the nature of the relationship. Other existing representations of medical information can be mapped directly into this model. Several authors have described the use of conceptual graph notation as a means of modeling descriptive findings, and the use of this notation is gaining popularity [14,15]. Conceptual graph notation maps onto relational data models and first-order predicate calculus, and is supported by the data model described in this report. Concepts or conceptual nodes in conceptual graph notation correspond to entity occurrences and conceptual relations map to attribute 'Relationship Between Medical Entity Occurrence #1 and Medical Entity Occurrence #2' of object MEDICAL ENTITY RELATIONSHIPS. The many-tomany relationship between MEDICAL ENTITY OCCURRENCES also supports the construction of polyhierarchies, and can therefore map to directed acyclic graphs [16] and semantic networks [17], which are other schemes used for representing information.

The process of defining, representing and managing all potential relationships in an electronic medical record using standard Entity-Relationship or Object-Oriented tools can result in a tremendous number of tables needing to maintained, and a pictorial view of the data model that is covered with hundreds to thousands of relationship lines and normalized relational tables. However, by defining a standard representation for the Unique_ID (or

relational 'primary key') for each medical object or entity, relationships between all entities can be defined using the high-level object-oriented model defined in this report. Limitations to this model include the inability to show mandatory relationships, such as that between a SERVICE and its corresponding ORDER. However, this model does not preclude explicit representation of known mandatory relationships.

In summary then, standards in medical informatics continue to gain importance as the need for reliable electronic medical data grows. A standardized unique identifier allows an electronic medical record to record important relationships between any two medical entities even though the relationships were not directly specified or known during data model design.

Acknowledgments

Special thanks to the Southern California Kaiser-Permanente Patient Care Data Modeling Team.

REFERENCES

- [1]. Institute of Medicine, Committee on Improving the Patient Record. The Computer-Based Patient Record: An Essential Technology for Health Care. Washington, DC: National Academy Press; 1991:190.
- [2]. American Hospital Association. Toward a National Health Information Infrastructure. Report of the Work Group on Computerization of Patient Records to the Secretary of the U.S. Department of Health and Human Services, April 1993.
- [3]. Clinton, WJ. American Health Security Act (Working Group Draft). The White House Domestic Policy Council, September 7, 1993.
- [4]. Barker R. <u>CASE Method Entity</u> <u>Relationship Modeling.</u> Reading, Mass: Addison-Wesley; 1990.
- [5]. Coad P, Yourdon E. <u>Object-oriented</u> <u>Analysis. 2nd Edition</u>. Prentice Hall; 1991.
- [6]. CEN (European Committee for Standardization)/ Technical Committee 251 -

- Medical Informatics. Project Team 009 Data interchange methodology for administrative and clinical data using intermittently connected devices. Interim Draft. CEN/TC251/PT009. January 12, 1994.
- [7]. HISPP/MSDS Joint Working Group for a Common Data Model, "Trial Use Standard for Medical Data Interchange--Information Model Methods" IEEE P1157.1, draft 1, 6/6/94.
- [8]. CEN (European Committee for Standardization)/ Technical Committee 251 Medical Informatics. Project Team 002 Terminology and coding systems of medical procedures. Interim Draft. CEN/TC251/PT002. November 10, 1993.
- [9]. Institute for Electrical and Electronic Engineers. Standard for medical device communications Overview and framework. IEEE P1073. Andover, MA; IEEE; 1992. [10]. Weed LL. Medical records that guide and
- teach. NEJM 1968; 278:593-600.
 [11]. Rozewski CM, Yahnke DP, Gottlieb MS,
- Hoffmann RG. A process for obtaining patient clinical information in the ambulatory setting. Computers and Biomedical Research 1993; 26(5):482-495.
- [12]. Ellwood PM. Shattuck lecture Outcomes management: A technology of patient experience. NEJM 1988; 318(23):1549-1556.
- [13]. Cote RA, Rothwell DJ, Beckett RS, Palotay JL (Eds). SNOMED International The systematized nomenclature of human and veterinary medicine. Northfield, Il; College of American Pathologists; 1993.
- [14]. Campbell KE, Das AK, Musen MA. A logical foundation for representation of clinical data. J Am Med Informatics Assoc 1994; 1(3): 218-232.
- [15]. Bernauer J. Conceptual graphs as an operational model for descriptive findings. SCAMC 1992; 214-218.
- [16]. De Vries W, Eidelman DH. Acyclic directed graphs for automatic image analysis of lung parenchymal geometry. Computers and Biomedical Research 1993; 26(4): 344-352. [17]. Lindberg DAB, Humphreys BL, McCray AT. The Unified Medical Language System. In: van Bemmel JH, McCray AT (Eds). Yearbook of Medical Informatics. The Netherlands; IMIA Publications; 1993, pg 41-51.